

# Universal Ideal Behavior and Macroscopic Work Relation of Linear Irreversible Stochastic Thermodynamics

Yi-An Ma<sup>\*</sup> and Hong Qian<sup>†</sup>

Department of Applied Mathematics  
University of Washington, Seattle  
WA 98195-3925, U.S.A.

March 29, 2015

## Abstract

We revisit the Ornstein-Uhlenbeck (OU) process as the fundamental mathematical description of linear irreversible phenomena, with fluctuations, near an equilibrium. By identifying the underlying circulating dynamics in a stationary process as the natural generalization of classical conservative mechanics, a bridge between a family of OU processes with equilibrium fluctuations and thermodynamics is established through the celebrated Helmholtz theorem. The Helmholtz theorem provides an emergent macroscopic “equation of state” of the entire system, which exhibits a universal ideal thermodynamic behavior. Fluctuating macroscopic quantities are studied from the stochastic thermodynamic point of view and a non-equilibrium work relation is obtained in the macroscopic picture, which may facilitate experimental study and application of the equalities due to Jarzynski, Crooks, and Hatano and Sasa.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Preliminaries</b>	<b>4</b>
2.1	Stationary Gaussian density and underlying conservative dynamics . . . . .	4
2.2	Zero energy reference: A hidden assumption in classical physics . . . . .	5

---

<sup>\*</sup>Email: yianma@u.washington.edu

<sup>†</sup>Email: hqian@u.washington.edu

<b>3</b>	<b>Free energy functions and functional</b>	<b>7</b>
3.1	Thermodynamic free energy functions $A(E, \alpha)$ . . . . .	7
3.2	Dynamic free energy functional $\Psi[f_\alpha(\mathbf{x}, t)]$ . . . . .	8
3.3	Universal equation of state of OU process . . . . .	9
<b>4</b>	<b>Circulating conservative flow and its invariant measures</b>	<b>10</b>
4.1	The conservative flow and general time reversal symmetries . . . . .	10
4.2	Conditional probability measure as invariant measure of the conservative flow . . . . .	11
4.3	Resolutions to the energy reference problem . . . . .	12
<b>5</b>	<b>Work equalities and fluctuation theorems</b>	<b>12</b>
5.1	The Jarzynski equality . . . . .	13
5.2	Crooks' approach . . . . .	14
5.2.1	Crooks' approach through adjoint processes . . . . .	15
5.3	Macroscopic work equalities . . . . .	15
<b>6</b>	<b>Discussion</b>	<b>16</b>
6.1	Absolute information theory and interpretive information theories . . . . .	17
<b>A</b>	<b>Derivation of work equalities</b>	<b>19</b>
A.1	The Jarzynski equality . . . . .	19
A.2	Crooks' approach . . . . .	21

# 1 Introduction

Gaussian fluctuation theory is one of the most successful branches of equilibrium statistical mechanics [1, 2]. Since the work of Onsager and Machlup [3, 4], the Ornstein-Uhlenbeck process (OUP) has become the stochastic, mathematical description of dynamic, linear irreversible phenomena [5]. It has been extensively discussed in the literature in the past [6, 7, 8, 9]. Several recent papers studied particularly the OUP without detailed balance [10, 11, 12]. In recent years, taking stochastic process rigorously developed by Kolmogorov as the mathematical representation, stochastic thermodynamics has emerged as the finite-time thermodynamic theory of mesoscopic systems, near and far from equilibrium [13, 14, 15, 16]. The fundamental aspects of this new development are the mathematical notion of stochastic entropy production [17, 18, 19], novel thermodynamic relationships collectively known as nonequilibrium work equalities, and fluctuation theorems [20, 21, 22, 23, 24], and the mathematical concept of non-equilibrium steady-state [25, 26, 27].

Fundamental to all these advances is the notion of *time reversal*. Newtonian dynamic equation, in Hamiltonian form:

$$\frac{dx_i}{dt} = \frac{\partial H(x_i, y_i)}{\partial y_i}, \quad \frac{dy_i}{dt} = -\frac{\partial H(x_i, y_i)}{\partial x_i}, \quad (1)$$

is a canonical example of dynamics with time-reversal symmetry [28]: Under transformation  $(t, x_i, y_i) \rightarrow (-t, x_i, -y_i)$ , Eq. (1) is invariant. This invariance requires that  $H(x_i, y_i) = H(x_i, -y_i)$ :  $H$  is usually a function of  $y_i^2$  and terms like  $\vec{B} \cdot \vec{y}$ , where  $\vec{B}$  changes sign upon time reversal such as a magnetic field with a Lorentz force. Adopting this definition to linear stochastic processes, one has a novel definition for *time reversibility* that is distinctly different from that of Kolmogorov's, as we shall show below.

Consider the linear stochastic differential equation

$$d\mathbf{X}(t) = -M(\alpha)\mathbf{X}(t)dt + \epsilon\Gamma d\mathbf{B}_t, \quad (2)$$

which is an OUP with parameters  $\alpha$  and  $\epsilon$ ;  $M$  and  $\Gamma$  are two  $n \times n$  constant matrices,  $\mathbf{B}_t$  is standard Brownian motion. We further assume that all the eigenvalues of  $M$  are strictly positive and  $\Gamma$  is non-singular. According to the concept of detailed balance, Eq. (2) can be uniquely written as [29, 30, 31, 32]

$$d\mathbf{X}(t) = -D\left\{\Xi^{-1} + \left(D^{-1}M - \Xi^{-1}\right)\right\}\mathbf{X}dt + \epsilon\Gamma d\mathbf{B}_t, \quad (3a)$$

where  $D$  and  $\Xi(\alpha)$  are positive definite matrices:  $D = \frac{1}{2}\Gamma\Gamma^T$  and  $M\Xi + \Xi M^T = 2D$ . If one identifies the two terms inside  $\{\dots\}$  as dissipative (transient) and conservative (perpetuate) motions, respectively, then a time reversible process should be defined as a statistical equivalence between the probability density of a finite path  $\{\mathbf{X}(t_0) = x_0, \mathbf{X}(t_1) = x_1, \dots, \mathbf{X}(t_n) = x_n\}$  in which  $t_0 < t_1 < \dots < t_n$ :

$$f(x_0, x_1, \dots, x_n),$$

and the probability density

$$f_{\mathbf{X}^\dagger(t_n)\mathbf{X}^\dagger(2t_n-t_{n-1})\dots\mathbf{X}^\dagger(2t_n-t_0)}(x_n, x_{n-1}, \dots, x_0)$$

in which the  $\mathbf{X}^\dagger(t)$  follows the adjoint stochastic differential equation [32, 33]

$$d\mathbf{X}^\dagger(t) = -D\left\{\Xi^{-1} - \left(D^{-1}M - \Xi^{-1}\right)\right\}\mathbf{X}^\dagger dt + \epsilon\Gamma d\mathbf{B}_t, \quad (3b)$$

with initial distribution for  $\mathbf{X}^\dagger(t_n)$  identical to that of  $\mathbf{X}(t_n)$ .

Recognizing the underlying circulating, conservative dynamics in Eqs. (3a) and (3b) allows us to connect a Hamiltonian structure with linear stochastic processes, and consequently develop a Helmholtz theorem, which historically has served as the fundamental

mathematical link between classical Newtonian mechanics and thermodynamics. For high dimensional stochastic processes, variables in the Helmholtz theorem provide the systems' underlying dynamics with a macroscopic picture. An ideal gas-like relation between a set of new, macroscopic variables emerges, confirming the simplicity of the OUP. A work-free energy equality in terms of the macroscopic thermodynamic variables, which are fluctuating with the underlying dynamics, captures the nature of the fluctuation in the underlying stochastic processes. We emphasize that even though the mathematical derivations are essentially the same, the physical meaning of the work relation is closer to the classical thermodynamics.

The paper is structured as follows. In Sec. 2, we first provide the necessary preliminaries on the OUP. Sec. 2.1 introduces the conservative dynamics as a part of the stationary behavior of the OUP. Sec. 2.2 then discusses a long neglected issue of zero energy reference. Secs. 3.1 and 3.2 introduces the stationary free energy function and the dynamic free energy functional. Sec. 3.3 studies the novel object of equation of state. It is shown that the OUP has a simply, universal ideal thermodynamic behavior. In Sec. 4, we turn to the circulating dynamics and its relation to classical mechanics as well as stochastic dynamics. Sec. 4.1 focuses on the simplicity of the circulating dynamics as being totally integrable. Sec. 4.2 contains a proof that the stationary probability density of OUP, conditioned on an invariant torus of the underlying conservative dynamics, analogous to a microcanonical ensemble, is an invariant measure of the latter. If the dynamics on an invariant torus is ergodic, then the conditional probability is the only, natural invariant measure on the torus. Work equalities and fluctuation theorems are discussed in Sec. 5. Using a macroscopic presentation of the Jarzynski equality, its relation to Helmholtz theorem is revealed in Sec. 5.3. The paper concludes with discussions in Sec. 6.

## 2 Preliminaries

### 2.1 Stationary Gaussian density and underlying conservative dynamics

The OUP in Eq. (3a) satisfies the important *fluctuation-dissipation relation*:  $2D\Xi^{-1} = \epsilon^2(\Gamma\Gamma^T) \times$  covariance matrix of the stationary OUP. In fact, it has a stationary Gaussian distribution  $Z^{-1}(\alpha)e^{-\varphi(\mathbf{x};\alpha)/\epsilon^2}$  in which  $Z(\alpha)$  is a normalization factor and  $\varphi(\mathbf{x};\alpha) = \mathbf{x}^T\Xi^{-1}(\alpha)\mathbf{x}$ . In addition, there is an underlying circulating dynamics

$$\frac{d\mathbf{x}}{dt} = -\left(M(\alpha) - D\Xi^{-1}(\alpha)\right)\mathbf{x}, \quad (4)$$

where the scalar  $\varphi(\mathbf{x};\alpha)$  is conserved [11]:

$$\frac{d}{dt}\varphi(\mathbf{x}(t);\alpha) = -2\mathbf{x}^T\Xi^{-1}\left(M - D\Xi^{-1}\right)\mathbf{x} = -\mathbf{x}^T\left(\Xi^{-1}M - M^T\Xi^{-1}\right)\mathbf{x} = 0. \quad (5)$$

In fact, this conservative dynamics can be expressed as [32]:

$$\frac{d\mathbf{x}}{dt} = -\frac{1}{2}\left(M\Xi - D\right)\nabla_{\mathbf{x}}\varphi(\mathbf{x}; \alpha), \quad (6)$$

where

$$\varphi(\mathbf{x}; \alpha) = \mathbf{x}^T \Xi^{-1}(\alpha) \mathbf{x}; \quad (7)$$

and  $(M(\alpha)\Xi - D)$  is skew-symmetric.

It is of paramount importance to recall that for a Markov process without detailed balance, its stationary dynamics is quantified by two mathematical objects: a stationary probability density and a stationary circulation [25, 34] characterized as a divergence-free, conservative vector field. In general, the latter accounts for the complexity arising from the system's dynamics [35]: how many integrals of motion does it have; whether the conservative dynamics is ergodic on an invariant set; etc. Many of the characteristics persist in the stationary stochastic process, and can be used to classify long time, complex behaviors in high dimensional systems. On the other hand, the dissipative (transient) dynamics plus noise drive the system towards the stationary distribution while characterizing “energy” fluctuations.

For the OUP in Eq. (3a), the conservative dynamics will be shown to be totally integrable. That is, symmetries would be implied through  $\lfloor n/2 \rfloor$  first integrals of motions, which are the natural generalizations of the time-reversal symmetries. The remaining part,  $d\mathbf{X}(t) = -\frac{1}{2}D\nabla_{\mathbf{x}}\varphi(\mathbf{x}; \alpha)dt + \epsilon\Gamma d\mathbf{B}_t$ , has a stationary dynamics that is detailed balanced.

It is worth noting that any  $\tilde{\varphi}(\mathbf{x}; \alpha) = \varphi(\mathbf{x}; \alpha) + C(\alpha)$  is also a valid substitution for the  $\varphi(\mathbf{x}; \alpha)$  in Eq. (7). As far as the stochastic dynamical system Eq. (3a) is concerned, there is no unique  $\tilde{\varphi}(\mathbf{x}; \alpha)$  as a function of both dynamic variable  $\mathbf{x}$  and parameter  $\alpha$ .

## 2.2 Zero energy reference: A hidden assumption in classical physics

The central object that connects classical Newtonian mechanics with equilibrium thermodynamics is the entropy function  $S(E, V, N)$ , with  $V$  and  $N$  being the volume and the number of particles of a classical mechanical system in a container, and  $E$  its total mechanical energy which is conserved according to Newton's Second Law of motion. In Hamilton's formulation Eq. (1),  $E$  is simply the initial value of the Hamiltonian function  $H(\{x_i\}, \{y_i\})$  in which  $x_i$  and  $y_i$  are the position and momentum of  $i$ th particle, respectively,  $1 \leq i \leq N$ .

We recognize that in the classical theory of mechanical motions, replacing  $H$  with  $\tilde{H}(\{x_i\}, \{y_i\}) = H(\{x_i\}, \{y_i\}) + C$ , where  $C$  is a constant, has absolutely no consequence to the mathematical theory. Therefore, with parameters contained in the Hamiltonian function, such as  $V$  and  $N$ ,  $H(\{x_i\}, \{y_i\}; V, N)$  and  $H(\{x_i\}, \{y_i\}; V, N) + C(V, N)$

are equivalent. In other words, classical mechanics only uniquely determines a Hamiltonian function up to an arbitrary function of all the non-dynamic parameters.

However, an additive function  $C(V, N)$  would cause non-uniqueness in the thermodynamic forces in the relation:

$$dS(E, V, N) = \left( \frac{\partial S}{\partial E} \right)_{V, N} [dE + p dV - \mu dN], \quad (8)$$

in which

$$p = \frac{\left( \frac{\partial S}{\partial V} \right)_{E, N}}{\left( \frac{\partial S}{\partial E} \right)_{V, N}} = - \left( \frac{\partial E}{\partial V} \right)_{S, N}, \quad \mu = \left( \frac{\partial E}{\partial N} \right)_{S, V}. \quad (9)$$

Corresponding to  $\tilde{H} = H + C(V, N)$  one has, for  $E$  as the initial values of  $H(\{x_i\}, \{y_i\}; V, N)$ , and  $\tilde{E}$  as the initial values of  $\tilde{H}(\{x_i\}, \{y_i\}; V, N)$ :

$$\tilde{p} = - \left( \frac{\partial \tilde{E}}{\partial V} \right)_{S, N} = p - \left( \frac{\partial C}{\partial V} \right)_N, \quad \tilde{\mu} = \mu + \left( \frac{\partial C}{\partial N} \right)_V. \quad (10)$$

Since pressure  $p$  has a mechanical interpretation, one can, by physical principle, uniquely determine the form of  $p$  as a function of  $V$ . The situation for  $\mu$  is much less clear: Since there is not an independent mechanical interpretation of the chemical potential other than the thermodynamic one given in Eq. (8), the non-uniqueness is inherent in the mathematical, as well as the physico-chemical theory. The problem has the same origin as *Gibbs' paradox* [37, 38].

In classical chemical thermodynamics, the Hamiltonian function as a function of varying number of particles  $N$ ,  $H(x_1, \dots, x_N, y_1, \dots, y_N) = \frac{1}{2} \sum_i^N m_i y_i^2 + V(x_1, \dots, x_N)$ , is uniquely determined via a Kirkwood charging process [36]:

$$V(x_1, \dots, x_N, x_{N+1}) = V(x_1, \dots, x_N) + \sum_{j=1}^N U_{j, N+1}(x_j, x_{N+1}), \quad (11)$$

where

$$\lim_{|x-y| \rightarrow \infty} U(x, y) = 0. \quad (12)$$

With this convention, the Hamiltonian for a molecular system is uniquely determined in chemical thermodynamics, which yields a consistent chemical potential  $\mu$ . How to generalize this chemical approach to Hamiltonian dynamics Eq. (1) with no clear separation between kinetic and potential parts, however, is unclear.

The problem of uniqueness of Hamiltonian function  $\tilde{H}$  is intimately related to the uniqueness of  $\tilde{\varphi}(\mathbf{x}; \alpha)$  in Sec. 2.1. As we shall show in the rest of this paper, the zero energy reference has deep implications to the theory of stochastic thermodynamics. The resolution to the problem will be discussed in Sec. 4.3.

### 3 Free energy functions and functional

As the notion of entropy, the definition of free energy is widely varied in the literature.<sup>1</sup> The most general features of free energy, perhaps, are: it is the difference between “internal energy” and entropy; it is the entropy under a “natural invariant measure”. In this section, we shall present two different types of free energies associated with the OU dynamics in Eq. (3a):

(i) Thermodynamic free energy of a stationary dynamics, as a function of mean internal energy  $E$  and parameter  $\alpha$ :  $A(E, \alpha)$ . We identify a “thermodynamic state” as a state of sustained motion, either for a deterministic conservative dynamics Eq. (4), or for a stochastic stationary process defined by Eq. (3a).

(ii) Dynamic free energy functional,  $\Psi[f(\mathbf{x}, t)]$ , for an instantaneous probability distribution  $f(\mathbf{x}, t)$ .

#### 3.1 Thermodynamic free energy functions $A(E, \alpha)$

With a particular given  $\varphi(\mathbf{x}; \alpha)$ , we now introduce two different free energy functions. The first one is defined following the microcanonical ensemble approach; definition of the second one follows Gibb’s canonical ensemble approach. While the second one is frequently being used in the work-free energy relation (discussed in Sec. 5), the two definitions agree perfectly in the large dimension limit.

The first thermodynamic free energy function,  $A_1(E, \alpha)$ , associated with the conservative deterministic motion of Eq. (4) on the surface of  $\varphi(\mathbf{x}; \alpha) = E$ , is obtained following the microcanonical ensemble approach through Boltzmann’s entropy function. Letting  $\sigma_B(E, \alpha)$  correspond to the entropy  $S$  and  $\Theta^{-1}(E, \alpha)$  correspond to  $(\frac{\partial S}{\partial E})$  in Eq. (8), we can define:

$$\begin{aligned}\sigma_B(E, \alpha) &= \ln \left( \int_{\varphi(\mathbf{x}; \alpha) \leq E} d\mathbf{x} \right) \\ &= \frac{n}{2} \ln E + \frac{1}{2} \ln \det \Xi(\alpha) + \ln V_n,\end{aligned}\tag{13a}$$

$$\Theta^{-1}(E, \alpha) = \left( \frac{\partial \sigma_B}{\partial E} \right)_\alpha = \frac{n}{2E},\tag{13b}$$

$$\begin{aligned}A_1(E, \alpha) &= E - \Theta \sigma_B \\ &= \frac{2E}{n} \left\{ -\frac{n}{2} \ln E - \frac{1}{2} \ln \det \Xi(\alpha) - \frac{n}{2} \ln(\pi) + \ln \Gamma \left( \frac{n}{2} + 1 \right) + \frac{n}{2} \right\},\end{aligned}\tag{13c}$$

---

<sup>1</sup>It has become increasingly clear that the Boltzmann’s entropy for a Hamiltonian dynamics is not unique: There are different geometric characterizations of the level sets of the Hamiltonian that can be acceptable choices. Neither is Shannon’s entropy in stochastic dynamics unique: other convex functions such as Tsallis’ entropy can also be found in the literature.

where  $V_n = \pi^{n/2} (\Gamma(\frac{n}{2} + 1))^{-1}$  is the volume of an  $n$ -dimensional Euclidean ball with radius 1.  $\Gamma(\cdot)$  is gamma function.  $n$  is the dimension of the OUP in Eq. (3a).

The second one,  $A_2(E, \alpha)$ , follows Gibbs' canonical ensemble approach via the "partition function"  $Z(\alpha)$ :

$$Z(\alpha) = \int_{\mathbb{R}^n} e^{-\varphi(\mathbf{x}; \alpha)/\epsilon^2} d\mathbf{x} = \left( (\pi\epsilon^2)^n \det \Xi(\alpha) \right)^{\frac{1}{2}}, \quad (14a)$$

$$\begin{aligned} A_2(\bar{E}, \alpha) &= -\epsilon^2 \ln Z(\alpha) \\ &= \frac{2\bar{E}}{n} \left\{ -\frac{n}{2} \ln \bar{E} - \frac{1}{2} \ln \det \Xi(\alpha) - \frac{n}{2} \ln(\pi) + \frac{n}{2} \ln \left( \frac{n}{2} \right) \right\}, \end{aligned} \quad (14b)$$

in which mean internal energy

$$\bar{E} = \frac{1}{Z(\alpha)} \int_{\mathbb{R}^n} \varphi(\mathbf{x}; \alpha) e^{-\varphi(\mathbf{x}; \alpha)/\epsilon^2} d\mathbf{x} = \frac{1}{2} n \epsilon^2. \quad (14c)$$

The two free energy functions  $A_1$  in Eq. (13c) and  $A_2$  in Eq. (14b) are different only by a function of  $n$  inside the  $\{\cdots\}$ . For large  $n$ ,  $\ln \Gamma(\frac{n}{2} + 1) \approx \frac{n}{2} \ln \left( \frac{n}{2} \right) - \frac{n}{2}$ . Therefore,  $A_1$  and  $A_2$  agree perfectly in the limit of  $n \rightarrow \infty$ .

### 3.2 Dynamic free energy functional $\Psi[f_\alpha(\mathbf{x}, t)]$

The thermodynamic free energy  $A_2(E, \alpha)$  in Sec. 3.1 sets a universal energy reference point for the entire family of stochastic dynamics in Eq. (3a) with different  $\alpha$ . For a given  $\alpha$ , the time-dependent probability density function  $f_\alpha(\mathbf{x}, t)$  follows the partial differential equation

$$\frac{\partial f_\alpha(\mathbf{x}, t)}{\partial t} = \nabla_{\mathbf{x}} \cdot \left( \epsilon^2 D \nabla_{\mathbf{x}} f_\alpha(\mathbf{x}, t) + M(\alpha) \mathbf{x} f_\alpha(\mathbf{x}, t) \right). \quad (15)$$

The  $f_\alpha(\mathbf{x}, t)$  represents an instantaneous "state" of the probabilistic system, which has a free energy functional

$$\begin{aligned} \Psi[f_\alpha(\mathbf{x}, t)] &= \int_{\mathbb{R}^n} \varphi(\mathbf{x}; \alpha) f_\alpha(\mathbf{x}, t) d\mathbf{x} - \left( -\epsilon^2 \int_{\mathbb{R}^n} f_\alpha(\mathbf{x}, t) \ln f_\alpha(\mathbf{x}, t) d\mathbf{x} \right) \\ &= \epsilon^2 \int_{\mathbb{R}^n} f_\alpha(\mathbf{x}, t) \ln \left( \frac{f_\alpha(\mathbf{x}, t)}{Z^{-1}(\alpha) e^{-\varphi(\mathbf{x}; \alpha)/\epsilon^2}} \right) d\mathbf{x} + A_2(\alpha). \end{aligned} \quad (16)$$

This is a dynamic generalization of the free energy functions in Sec. 3.1. It has two important properties. First,

$$\lim_{t \rightarrow \infty} \Psi[f_\alpha(\mathbf{x}, t)] = A_2(\alpha). \quad (17)$$

Second [32, 33],

$$\frac{d}{dt} \Psi[f_\alpha(\mathbf{x}, t)] \leq 0, \quad (18)$$



in which the equality holds if and only if  $f_\alpha(\mathbf{x}, t)$  reaches its stationary distribution  $Z^{-1}(\alpha)e^{-\varphi(\mathbf{x};\alpha)/\epsilon^2}$ . The negated rate of change in the dynamic free energy functional,  $-\mathrm{d}\Psi/\mathrm{d}t$ , is widely recognized as non-adiabatic entropy production rate.

The entropy production rate also has a finite time, stochastic counterpart in terms of the logarithm of the likelihood ratio:

$$-\frac{\mathrm{d}}{\mathrm{d}t}\Psi[f_\alpha(\mathbf{x}, t)] = \lim_{s \rightarrow t} E^\mathbb{P} \left[ \frac{1}{|s-t|} \ln \left( \frac{f(\mathbf{X}(\tau)|t \leq \tau \leq s)}{f_{\mathbf{X}^\dagger}(\hat{\mathbf{X}}(\tau)|t \leq \tau \leq s)} \right) \right], \quad (19)$$

where  $\hat{\mathbf{X}}(\tau) = \mathbf{X}(s - \tau + t)$ , and the expectation  $E^\mathbb{P}[\dots]$  is carried out over the diffusion process defined by Eq. (3a) and the corresponding Eq. (15):

$$f(\mathbf{X}(\tau)|t \leq \tau \leq s) \propto \exp \left[ -2 \int_t^s \left( \dot{\mathbf{X}}(\tau) + M\mathbf{X}(\tau) \right)^T D \left( \dot{\mathbf{X}}(\tau) + M\mathbf{X}(\tau) \right) \mathrm{d}\tau \right]. \quad (20)$$

### 3.3 Universal equation of state of OU process

With the introduction of the internal energy  $E$  and the parameter  $\alpha$ , the thermodynamic relation Eq. (8) - known as the Helmholtz theorem - for the OUP model can be expressed by  $\sigma_B$ ,  $\alpha$  and their conjugate variables. We notice that  $\alpha$  enters Eq. (13) only through  $\det \Xi(\alpha)$ . If one measures  $\alpha$  through  $\tilde{\alpha} = \det \Xi(\alpha)$ , then the Helmholtz theorem writes:

$$\begin{aligned} \mathrm{d}E &= \Theta(E, \tilde{\alpha}) \mathrm{d}\sigma_B - F_{\tilde{\alpha}}(E, \tilde{\alpha}) \mathrm{d}\tilde{\alpha} \\ &= \left( \frac{\partial \sigma_B}{\partial E} \right)^{-1} \mathrm{d}\sigma - \left( \frac{\partial \sigma_B}{\partial \tilde{\alpha}} \right) \left( \frac{\partial \sigma_B}{\partial E} \right)^{-1} \mathrm{d}\tilde{\alpha}. \end{aligned} \quad (21)$$

The two conjugate variables,  $\Theta$  and  $F_{\tilde{\alpha}}$ , correspond to the macroscopic quantities in classical thermodynamics as temperature and force.<sup>2</sup>

Following either Boltzmann's microcanonical or Gibbs' canonical approach, Sec. 3.1 revealed that  $E = \frac{1}{2}n\Theta$  in which  $\theta = \frac{1}{2n}\Theta$  could be interpreted as an "absolute temperature". Since the absolute temperature  $\theta$  is a fluctuating quantity with respect to  $E$  and  $\tilde{\alpha}$ , it may, in general, not bear a simple relationship with the noise strength  $\epsilon^2$ . But here in OUP, by comparing the microcanonical approach with the canonical one, we note that the mean absolute temperature  $\bar{\theta} = \frac{\epsilon^2}{2n}$ .

The thermodynamic conjugate variable of  $\tilde{\alpha}$ , the  $\tilde{\alpha}$ -force:

$$F_{\tilde{\alpha}} = \Theta \left( \frac{\partial \sigma_B(E, \tilde{\alpha})}{\partial \tilde{\alpha}} \right)_E = \frac{n\theta}{\tilde{\alpha}}. \quad (22)$$

A mathematical relation between  $\tilde{\alpha}$ ,  $F_{\tilde{\alpha}}$ , and  $\theta$  is called an *equation of state* in classical thermodynamics.

---

<sup>2</sup>The force here should be understood as Onsager's thermodynamic force: corresponding to a spatial displacement is a mechanical force; to a change in number of particles is Gibbs' chemical potential; to a variation in a parameter through a Maxwell demon then is an informatic force [39, 40].

The “internal energy”  $E$  being a sole function of temperature  $\theta$ , and the product of thermodynamic conjugate variables,  $\tilde{\alpha}F_{\alpha}^{-}$ , equaling to  $n\theta$ , are hallmarks of thermodynamic behavior of ideal gas and ideal solution. We thus conclude that the OUP has a universal ideal thermodynamic behavior.

## 4 Circulating conservative flow and its invariant measures

After discussing the energy function and stationary probability, we now focus on the dynamic complexity of the system and study the circulating, conservative dynamics. The universal ideal thermodynamic behavior reveals one aspect of the simplicity in OUP; another is reflected in the divergence-free motions. For the linear conservative dynamics, Eq. (4), its structure is known to be simple: the vector field is integrable.

The conservative dynamics in Eq. (4) can be proved to be purely cyclic (e.g., periodic, or quasi-periodic on an invariant torus). Because the skew-symmetric matrix  $(M - D\Xi^{-1})$  has only pairs of imaginary eigenvalues  $\{\lambda_{\ell}|1 \leq \ell \leq n\}$ . We can also find real Jordan form of  $(M - D\Xi^{-1})$ :  $QJQ^{-1}$ , where  $J$  is block diagonal, with  $2 \times 2$  skew-symmetric blocks:

$$\text{Im} [\lambda_{(2i-1)}] \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

being the  $i$ th block on the diagonal. Natural coordinates for the conservative flow Eq. (4) is therefore:  $\mathbf{y} = Q^{-1}\mathbf{x}$ .

### 4.1 The conservative flow and general time reversal symmetries

Poisson bracket  $\{\cdot, \cdot\}$  can be defined for the linear conservative system as:  $\{\varphi(\mathbf{x}), \psi(\mathbf{x})\} = \nabla\varphi(\mathbf{x})^T(M - D\Xi^{-1})\nabla\psi(\mathbf{x})$ . Then the conservative flow expressed in terms of its Hamiltonian function  $\varphi(\mathbf{x})$  is:

$$\dot{x}_i = \left\{x_i, \frac{1}{2}\varphi(\mathbf{x})\right\}. \quad (23)$$

First integrals  $I_i$  of the conservative flow are:

$$I_i = y_{2i-1}^2 + y_{2i}^2 = \mathbf{x}^T Q^{-T} I_{(2i-1) \sim (2i)} Q^{-1} \mathbf{x}, \quad 1 \leq i \leq \left\lfloor \frac{n}{2} \right\rfloor. \quad (24)$$

Here,  $I_{(2i-1) \sim (2i)}$  denotes the diagonal matrix with 1 on  $(2i-1)$ -th to  $(2i)$ -th diagonal entries, and zero everywhere else.

The conservative flow is totally integrable, and can be written in canonical action-angle variables. Angular coordinates  $\theta_i$  accompanying  $I_i$  can be found as:

$$\theta_i = \text{Im} [\lambda_{2i-1}]^{-1} \cdot \arctan \left( \frac{y_{2i-1}}{y_{2i}} \right), \quad 1 \leq i \leq \left\lfloor \frac{n}{2} \right\rfloor. \quad (25)$$

Hence, in the canonical action-angle variables,  $\varphi = \sum_{i=1}^{\lfloor n/2 \rfloor} I_i$ ,

$$\begin{cases} \theta'_i = \frac{\partial \varphi}{\partial I_i} = 1 \\ I'_i = -\frac{\partial \varphi}{\partial \theta_i} = 0. \end{cases} \quad (26)$$

There are  $\lfloor \frac{n}{2} \rfloor$  first integrals, but for the given Poisson bracket, one combination of them is unique, which is the Hamiltonian  $\varphi$  that connects to the stationary distribution and generates the conservative flow.

In the action-angle variables, it is observable that the system bears the following symmetries:  $(t, \theta_i, I_i) \longrightarrow (-t, -\theta_i, (-1)^{k_i} I_i)$ , where  $\{k_i\}$  is a sequence of 0 and 1. Taking  $\{k_i\}$  as a sequence of zeros, we recover the time-reversal invariance in Eq. (1). Hence, for general  $\{k_i\}$ , those symmetries are the natural generalizations of the time-reversal symmetry, as displayed in classical Hamiltonian systems.

## 4.2 Conditional probability measure as invariant measure of the conservative flow

The OUP yields an equilibrium probability density function for  $\mathbf{X}$ :  $f_{\mathbf{X}}^{eq}(\mathbf{x}; \alpha) = Z^{-1}(\alpha) e^{-\varphi(\mathbf{x}; \alpha)/\epsilon^2}$ . In this section, we calculate the *conditional probability density* for  $\mathbf{X}$  restricted on an equal energy surface  $\mathfrak{D}_{\varphi=E} = \{\mathbf{x} | \mathbf{x} \in \mathbb{R}^n, \varphi(\mathbf{x}; \alpha) = E\}$  and prove it to be an invariant measure of the conservative dynamics, Eq. (4), restricted on  $\mathfrak{D}_{\varphi=E}$ . Therefore, in the absence of fluctuation and dissipation, our definition of “equilibrium free energy” (Eq. (13c)) in stochastic thermodynamics retreats to the Boltzmann’s microcanonical ensemble approach in classical mechanics.

One can obtain a conditional probability density for  $\mathbf{X}$  restricted on an equal energy surface  $\mathfrak{D}_{\varphi=E} = \{\mathbf{x} | \mathbf{x} \in \mathbb{R}^n, \varphi(\mathbf{x}; \alpha) = E\}$  as:

$$Z(\alpha) = \int_{\mathbb{R}^n} e^{-\varphi(\mathbf{x}; \alpha)/\epsilon^2} d\mathbf{x} = \int_{\varphi_{min}(\alpha)}^{\infty} \exp\left(-\frac{E}{\epsilon^2} + S(E, \alpha)\right) dE, \quad (27)$$

in which [41]

$$S(E, \alpha) = \ln \left( \frac{\partial}{\partial E} \int_{\varphi(\mathbf{x}; \alpha) \leq E} d\mathbf{x} \right)_{\alpha} = \ln \left( \oint_{\varphi(\mathbf{x}; \alpha) = E} \frac{d\Sigma^{n-1}}{\|\nabla_{\mathbf{x}} \varphi(\mathbf{x}; \alpha)\|} \right). \quad (28)$$

The conditional probability density at  $\mathbf{x} \in \mathfrak{D}_{\varphi=E}$  is:

$$\frac{e^{-S(E, \alpha)}}{\|\nabla_{\mathbf{x}} \varphi(\mathbf{x}; \alpha)\|} = \frac{1}{\frac{n}{2} V_n E^{\frac{n}{2}-1} (\det \Xi)^{-\frac{1}{2}} \|\Xi^{-1} \mathbf{x}\|}. \quad (29)$$

Note this conditional probability is one of the invariant measures of the conservative dynamics Eq. (4) restricted in  $\mathfrak{D}_{\varphi=E}$ .

To prove this fact, define the dynamics of the conservative part as:  $\mathbf{S}_t$ , mapping a measurable set  $A \rightarrow \mathbf{S}_t(A)$ . Then measure of a set  $A \subseteq \mathcal{D}_{\varphi=E}$  under  $d\mu = e^{-S(E,\alpha)} \|\nabla_{\mathbf{x}}\varphi(\mathbf{x}; \alpha)\|^{-1} d\Sigma^{n-1}$  is:

$$\begin{aligned} \int_A \frac{e^{-S(E,\alpha)}}{\|\nabla_{\mathbf{x}}\varphi(\mathbf{x}; \alpha)\|} d\Sigma^{n-1} &= e^{-S(E,\alpha)} \int_A \delta(E - \varphi(\mathbf{x})) d\mathbf{x} \\ &= e^{-S(E,\alpha)} \int_{\mathbf{S}_t^{-1}(A)} \delta(E - \varphi(\mathbf{x})) d\mathbf{x} \\ &= \int_{\mathbf{S}_t^{-1}(A)} \frac{e^{-S(E,\alpha)}}{\|\nabla_{\mathbf{x}}\varphi(\mathbf{x}; \alpha)\|} d\Sigma^{n-1}, \end{aligned} \quad (30)$$

since  $\mathbf{S}_t^{-1}(A) \subseteq \mathcal{D}_{\varphi=E}$  and  $\mathbf{S}_t$  is volume preserving. In general, if the dynamics is ergodic on the entire  $\mathcal{D}_{\varphi=E}$ , then its invariant measure  $\mu$  is the *physical measure*:  $\mu$ -average equals time average along a trajectory; if there are other first integrals for the conservative dynamics, then  $\mu$  can be projected further to lower dimensional invariant sets.

### 4.3 Resolutions to the energy reference problem

Up to now, there are clearly several possibilities to uniquely determine the free additive function  $C(V, N)$  in the Hamiltonian  $H(\{x_i\}, \{y_i\}; V, N)$  discussed in Introduction.

- (1)  $C(V, N)$  is chosen such that the global minimum of  $H = 0$  for each and every  $V$  and  $N$ . This is widely used, implicitly, in application practices, as in our Eq. (7).
- (2)  $C(V, N)$  is chosen according to the “equilibrium free energy”:

$$-\epsilon^2 \ln \int e^{-E/\epsilon^2 + S(E,V,N)} dE = 0. \quad (31)$$

Note that this is precisely the “energy function” in Hatano and Sasa [24].

- (3) Extra information concerning the fluctuations in  $V$ , such as in an isobaric ensemble, and fluctuations in  $N$  in grand ensemble, provides an empirically determined basis for the free energy scale.

In terms of the theory of probability, choice (2) uniquely determines the energy reference point according to a conditional probability, and in choice (3) it is uniquely determined according to a marginal distribution. How to normalize a probability, which has always been considered non-consequential in statistical physics, seems to be a fundamental problem in the physics of complex systems.

## 5 Work equalities and fluctuation theorems

The previous discussions suggest that while a great deal of complexity of a detailed, mesoscopic stationary dynamics is captured by the circulating conservative dynamics, OUP also has a macroscopic state of motion that is defined by the internal energy  $E$ , or equivalently

the level sets of  $\varphi(\mathbf{x}), \mathcal{D}_{\varphi=E} \subset \mathbb{R}^n$ . Thus, from the macroscopic point of view, a stochastic system could be studied through the one-dimensional (1-D) time sequence of fluctuating internal energy  $E$ , as a function of  $t$ , or the change in  $E$  due to changes in the parameter  $\alpha$ .

The celebrated Jarzynski equality connects the mesoscopic fluctuating force with the change in free energy. We present this result through a projection from  $n$ -D phase space to 1-D function  $E(t)$  that facilitates experimental verification of the work-free energy relation. This approach reveals a close connection between the Jarzynski equality and the Helmholtz theorem. We start with stating the Jarzynski and Crooks' equalities, with the mathematical proofs collected in the Appendix A for readers' convenience. We then demonstrate the novel formulation of the Jarzynski equality in the projected space.

We have shown that  $A_2(\alpha)$  is uniquely determined only up to a particular  $\varphi(\mathbf{x}; \alpha)$ . As shown below, the existence of an  $A_2(\alpha)$  has a paramount importance in the theories of work equalities, in which the notion of a common energy for a family of stochastic dynamical systems with different  $\alpha$  has to be given *a priori* [42].

## 5.1 The Jarzynski equality

The macroscopic  $\alpha$ -force in the Helmholtz theorem, as a function of  $E$  and  $\alpha$ , is defined through Boltzmann's entropy  $\sigma_B$ . The Jarzynski equality, on the other hand, concerns with a mesoscopic  $\alpha$ -force,

$$F_\alpha(\mathbf{x}; \alpha) = - \left( \frac{\partial \varphi(\mathbf{x}; \alpha)}{\partial \alpha} \right)_{\mathbf{x}}, \quad (32)$$

and the statistical behavior of its corresponding *stochastic work*

$$W[\mathbf{X}(\tau), \alpha(\tau)] = \int_0^t F_\alpha(\mathbf{X}(\tau); \alpha(\tau)) \left( \frac{d\alpha(\tau)}{d\tau} \right) d\tau. \quad (33)$$

The Jarzynski equality dictates that if the initial distribution of  $\mathbf{X}(\tau)$  follows the equilibrium distribution, then [20]

$$\left\langle e^{-\frac{1}{\epsilon^2} W[\mathbf{X}(\tau), \alpha(\tau)]} \right\rangle_{[\mathbf{x}(\tau), \alpha(\tau)]} = e^{-\frac{1}{\epsilon^2} \Delta A_2(\alpha)}, \quad (34)$$

where the average of a functional over the ensemble of paths is defined as:

$$\left\langle G[\mathbf{X}(\tau), \alpha(\tau)] \right\rangle_{[\mathbf{x}(\tau), \alpha(\tau)]} = \int G[\mathbf{X}(\tau), \alpha(\tau)] \mathcal{P}[\mathbf{X}(\tau), \alpha(\tau)] \mathcal{D}[\mathbf{X}(\tau)], \quad (35)$$

in which  $\mathcal{P}[\mathbf{X}(\tau), \alpha(\tau)] \mathcal{D}[\mathbf{X}(\tau)]$ , is an infinite-dimensional probability distribution for the entire paths  $[\mathbf{X}(\tau)]$ .

It is clear from the proof in Appendix A that the Jarzynski equality is general for Markov processes with or without detailed balance. Several recent papers have studied extensively the latter case [43, 44].

## 5.2 Crooks' approach

G. E. Crooks' approach, when applied to processes without detailed balance [24], considers the probability functional of a backward path  $\mathcal{P}[\check{\mathbf{X}}(t)|\check{\mathbf{X}}(0);\check{\alpha}(t)]$  over a forward one  $\mathcal{P}[\mathbf{X}(t)|\mathbf{X}(0);\alpha(t)]$ , where both the initial and final distribution of  $\mathbf{X}(\tau)$  follows the equilibrium distribution:

$$\begin{aligned}\frac{\mathcal{P}[\check{\mathbf{X}}(\tau), \check{\alpha}(\tau)]}{\mathcal{P}[\mathbf{X}(\tau), \alpha(\tau)]} &= \exp \left( \frac{Q[\mathbf{X}(\tau), \alpha(\tau)] - Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)] - \Delta\varphi + \Delta A_2}{\epsilon^2} \right) \\ &= \exp \left( \frac{-W[\mathbf{X}(\tau), \alpha(\tau)] - Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)] + \Delta A_2}{\epsilon^2} \right),\end{aligned}\quad (36)$$

where  $\check{\mathbf{X}}(\tau) = \mathbf{X}(t - \tau)$ ,  $\check{\alpha}(\tau) = \alpha(t - \tau)$ ; and

$$\begin{aligned}Q[\mathbf{X}(\tau), \alpha(\tau)] &= \int_0^t \frac{\partial \varphi}{\partial \mathbf{X}} \dot{\mathbf{X}} \, d\tau, \\ Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)] &= -2 \int_0^t \dot{\mathbf{X}}^T D(\alpha)^{-1} (M(\alpha) - D(\alpha) \Xi^{-1}(\alpha)) \mathbf{X} \, d\tau\end{aligned}\quad (37)$$

are the heat dissipation and the house-keeping heat respectively.

If a process describes a physical system in equilibrium, which is expected to be “microscopic reversible” in [22], then

$$\left\langle \frac{\mathcal{P}[\check{\mathbf{X}}(\tau), \check{\alpha}(\tau)]}{\mathcal{P}[\mathbf{X}(\tau), \alpha(\tau)]} \right\rangle_{[\mathbf{X}(\tau), \alpha(\tau)]} = \int \mathcal{D}[\mathbf{X}(\tau), \alpha(\tau)] \mathcal{P}[\check{\mathbf{X}}(\tau), \check{\alpha}(\tau)] = 1. \quad (38)$$

On the other hand, if the system is in detailed balance for each and every  $\alpha$ ,  $M(\alpha) - D(\alpha) \Xi^{-1}(\alpha) = 0$ , then the house-keeping heat  $Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)] \equiv 0$ . Therefore, path-ensemble average of Eq. (36) gives:

$$\begin{aligned}1 &= e^{\Delta A_2/\epsilon^2} \left\langle e^{-W[\mathbf{X}(\tau), \alpha(\tau)]/\epsilon^2 - Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)]/\epsilon^2} \right\rangle_{[\mathbf{X}(\tau), \alpha(\tau)]} \\ &= e^{\Delta A_2/\epsilon^2} \left\langle e^{-W[\mathbf{X}(\tau), \alpha(\tau)]/\epsilon^2} \right\rangle_{[\mathbf{X}(\tau), \alpha(\tau)]}.\end{aligned}\quad (39)$$

This is Hatano-Sasa's result [24]. For systems without detailed balance,  $Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)]$  measures the magnitude of the divergence-free vector field, or the extent to which the system is away from detailed balance, even when stationary distribution is attained. At the same time,

$$\left\langle \frac{\mathcal{P}[\check{\mathbf{X}}(\tau), \check{\alpha}(\tau)]}{\mathcal{P}[\mathbf{X}(\tau), \alpha(\tau)]} \right\rangle_{[\mathbf{X}(\tau), \alpha(\tau)]}$$

measures how much on average the behavior of backward paths is statistically different from forward ones.

### 5.2.1 Crooks' approach through adjoint processes

Jarzynski's approach is based on a mesoscopic  $\alpha$ -force; while Crooks' approach concerns with the stochastic entropy production rate which reflects "heat dissipation". Therefore, for systems with detailed balance, they are essentially the same result according to the First Law of thermodynamics. For systems without detailed balance, one can again obtained a Jarzynski-like equality from the probability  $P$  of the forward path over the adjoint probability  $P^\dagger$  of the backward one, according to the notion of *time reversal* in Eq. (3b):

$$P^\dagger(\mathbf{X}(\tau)|\mathbf{X}(\tau + d\tau); \alpha(\tau)) = P(\mathbf{X}(\tau + d\tau)|\mathbf{X}(\tau); \alpha(\tau)) \times \left( \frac{f_{\mathbf{X}}^{eq}(\mathbf{X}(\tau); \alpha(\tau))}{f_{\mathbf{X}}^{eq}(\mathbf{X}(\tau + d\tau); \alpha(\tau))} \right). \quad (40)$$

Thus, whether a system is in detailed balance or not, one has the Hatano-Sasa equality [24]:

$$1 = \left\langle \frac{\mathcal{P}[\mathbf{X}(\tau), \alpha(\tau)]}{\mathcal{P}^\dagger[\tilde{\mathbf{X}}(\tau), \tilde{\alpha}(\tau)]} \right\rangle_{[\mathbf{x}(\tau), \alpha(\tau)]} = e^{\Delta A_2/\epsilon^2} \left\langle e^{-W[\mathbf{X}(\tau), \alpha(\tau)]/\epsilon^2} \right\rangle_{[\mathbf{x}(\tau), \alpha(\tau)]}. \quad (41)$$

### 5.3 Macroscopic work equalities

We are now in the position to study the work-free energy relation from a macroscopic view. Essentially, we will consider the stochastic, fluctuating  $(E(t), \tilde{\alpha}(t))$  instead of  $(\mathbf{X}(t); \alpha(t))$  directly. In doing so, we are observing the evolution in the probability distribution of  $E$  through a projection from  $(\mathbf{X}; \alpha)$  to  $(E, \tilde{\alpha})$ . With the projection of the  $n$ -dimensional phase space to the one-dimensional time series  $E(t)$ , the stationary probability density function  $f_{\mathbf{E}}^{ss}(E, \tilde{\alpha})$  of  $E$  with  $\tilde{\alpha}$  is also a projection of the original stationary probability density function  $f_{\mathbf{X}}^{ss}(\mathbf{x}; \tilde{\alpha})$  in Euclidean space (as discussed in Sec. 4.2):

$$f_{\mathbf{E}}^{ss}(E, \tilde{\alpha}) = \oint_{\varphi(\mathbf{x}; \tilde{\alpha})=E} \frac{f_{\mathbf{X}}^{ss}(\mathbf{x}; \tilde{\alpha}) d\Sigma^{n-1}}{\|\nabla_{\mathbf{x}} \varphi(\mathbf{x}; \tilde{\alpha})\|} = \frac{1}{Z(\tilde{\alpha})} \exp \left( -\frac{E}{\epsilon^2} + S(E, \tilde{\alpha}) \right), \quad (42)$$

in which

$$S(E, \tilde{\alpha}) = \ln \left( \oint_{\varphi(\mathbf{x}; \tilde{\alpha})=E} \frac{d\Sigma^{n-1}}{\|\nabla_{\mathbf{x}} \varphi(\mathbf{x}; \tilde{\alpha})\|} \right). \quad (43)$$

For the process of  $(E(t), \tilde{\alpha}(t))$ , the total internal energy is no longer  $E$  itself. But rather, it would include the "entropic effect",  $S(E, \tilde{\alpha})$ , caused by the curved space structure, and become  $\mathcal{A}(E, \tilde{\alpha})$ , as will be discussed in more detail in [45]:

$$\mathcal{A}(E, \tilde{\alpha}) = E - \epsilon^2 S(E, \tilde{\alpha}). \quad (44)$$

The Helmholtz theorem for the new  $(E(t), \tilde{\alpha}(t))$  process reads:

$$\begin{aligned} d\mathcal{A} &= \tilde{\Theta}(E, \tilde{\alpha})d\sigma - \tilde{F}_{\tilde{\alpha}}(E, \tilde{\alpha})d\tilde{\alpha} \\ &= \left( \left( \frac{\partial \sigma}{\partial E} \right)^{-1} - \epsilon^2 \frac{\partial S}{\partial \sigma} \right) d\sigma - \left( \left( \frac{\partial \sigma}{\partial \tilde{\alpha}} \right) \left( \frac{\partial \sigma}{\partial E} \right)^{-1} + \epsilon^2 \frac{\partial S}{\partial \tilde{\alpha}} \right) d\tilde{\alpha}. \end{aligned} \quad (45)$$

Hence, the total force that does the work in this new coordinate is:

$$\tilde{F}_{\tilde{\alpha}}(E, \tilde{\alpha}) = -\frac{\partial \mathcal{A}(E, \tilde{\alpha})}{\partial \tilde{\alpha}} = \frac{1}{n} \cdot \frac{E}{\tilde{\alpha}} + \epsilon^2 \frac{\partial S(E, \tilde{\alpha})}{\partial \tilde{\alpha}}, \quad (46)$$

where  $\epsilon^2(\partial S(E, \tilde{\alpha})/\partial \tilde{\alpha})$  is what chemists called an “entropic force”.

Now we define the work that external environment has done to the system through the controlled change of  $\tilde{\alpha}(t)$  as:

$$W[E(\tau), \tilde{\alpha}(\tau)] = -\int_0^t \tilde{F}_{\tilde{\alpha}}(E, \tilde{\alpha}) \dot{\tilde{\alpha}} d\tau. \quad (47)$$

Then the work-free energy relation in macroscopic variables is:

$$\left\langle e^{-W[E(\tau), \tilde{\alpha}(\tau)]} \right\rangle_{[E(\tau), \tilde{\alpha}(\tau)]} = \frac{Z(\tilde{\alpha}(t))}{Z(\tilde{\alpha}(0))} = \exp \left( -\frac{\Delta A_2(\alpha)}{\epsilon^2} \right). \quad (48)$$

Therefore, the averaged minus exponential of work is equal to the minus exponential of free energy difference. Here, we notice that the free energy stays the same through the change of free variables, as a result of Eq. (27):

$$Z(\alpha) = \int_{\mathbb{R}^n} e^{-\varphi(\mathbf{x}; \alpha)/\epsilon^2} d\mathbf{x} = \int_{\varphi_{min}(\alpha)}^{\infty} \exp \left( -\frac{E}{\epsilon^2} + S(E, \alpha) \right) dE.$$

## 6 Discussion

In the present work, using the OUP as an example, we have illustrated a possible method of deriving emergent, macroscopic descriptions of a complex stochastic dynamics from its mesoscopic *law of motion*. In recent years, there is a growing awareness of the role of probabilistic reasoning as *the* logic of science [46, 47]. In this framework, prior information, data, and probabilistic deduction are three pillars of a scientific theory. In fields with very complex dynamics, statistical inferences focus on the latter two aspects starting with data. In physical sciences that includes chemistry, and cellular biology, the prior plays a fundamental role as a feasible “mechanism” which enters a scientific model based on “established knowledge” — no biochemical phenomena should violate the physical laws of mechanics and thermodynamics. Indeed, many priors have been rigorously formalized in terms of mathematical theories. Unfortunately, most of these theories are expressed in terms of deterministic mathematics for very simple individual “particles”; obtaining a meaningful



probabilistic prior for a realistic, macroscopic-level system requires a computational task that is neither feasible nor meaningful [52, 53]. Nonlinear stochastic dynamical study is the mathematical deductive process that formulates probabilistic prior based on a given mechanism.

Open systems, when represented in terms of Markov processes, are ubiquitously non-symmetric processes according to Kolmogorov’s terminology. This is one of the lessons we learned from the open-chemical systems theory. The non-symmetry can be quantified by *entropy production* [25]. For discrete-state Markov processes, symmetric processes are equivalent to Kolmogorov’s cycle condition [54]. Interestingly, concepts such as cycle condition, detailed balance, dissipation and irreversible entropy production had all been independently discovered in chemistry: Wegscheider’s relation in 1901 [55], detailed balance by G.N. Lewis in 1925 [56], Onsager’s dissipation function in 1931 [57], and the formulation of entropy production in the 1940s [58, 59].

A non-symmetric Markov process implies circulating dynamics in phase space. Such dynamics is not necessarily dissipative, as exemplified by harmonic oscillators in classical mechanics. One of us has recently pointed out the important distinction between *overdamped thermodynamics* and *underdamped thermodynamics* [32]. The present paper is a study of OUP in terms of the latter perspective, in which we have identified the unbalanced circulation as a conservative dynamics, a hallmark of the generalized underdamped thermodynamics [33]. In terms of this conservative dynamics, Boltzmann’s entropy function naturally enters stochastic thermodynamics, and we discover a relation between the Helmholtz theorem [60] and the various work relations.

In the past, studies on stochastic thermodynamics with underdamped mechanical motions have always required an explicit identification of even and odd variables. See recent [61] and [32] and the references cited within. One of us has introduced a more general stochastic formulation of “underdamped” dynamics, with thermodynamics, in which circulating motion can be a part of a conservative motion [32] without dissipation. The present work is an in-depth study of the OUP within this new framework. It seems to us that even the term “nonequilibrium” in the literature has two rather different meanings: From a classical mechanical standpoint, any system with a stationary current is “nonequilibrium”, even though it can be non-dissipative. From a statistical mechanics stand point, on the other hand, “nonequilibrium”, “irreversible”, and “dissipative” are almost all synonymous.

## 6.1 Absolute information theory and interpretive information theories

We now discuss two rather different perspectives on the nature of information theory, or theories [62, 46].

First, in the framework of classical physics in terms of Newtonian mechanics, Boltz-

mann's law, and Gibbs' theory of chemical potential, there is a universal First Law of Thermodynamics based on the function  $S(E, V, N, \alpha)$  where  $S$  is the Boltzmann's entropy of a conservative dynamical system at total energy  $E$ , e.g., Hamiltonian  $H(\{x_i\}, \{y_i\}) = E$ , with  $V$  and  $N = (n_1, n_2, \dots, n_m)$  being the volume and numbers of particles in the chemomechanical system, and  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_\nu)$  represents controllable parameters of the system. Then one has

$$dE = \left( \frac{\partial S}{\partial E} \right)_{V, N, \alpha}^{-1} dS - p dV + \mu dN - \left( \frac{\partial S}{\partial E} \right)_{V, N, \alpha}^{-1} \left( \frac{\partial S}{\partial \alpha} \right)_{E, V, N} d\alpha, \quad (49)$$

in which  $(\partial S / \partial E)_{V, N, \alpha}^{-1}$  is absolute temperature.  $p$  and  $\mu$  are pressure and chemical potential, they are the corresponding thermodynamic forces for changing volume  $V$  and number of particles  $N$ , respectively. It is natural to suggest that if an agent is able to manipulate a classical system through changing  $\alpha$  while holding  $S$ ,  $V$ , and  $N$  constant, then he or she is providing to, or extracting from, the classical system *non-mechanical, non-chemical work*. It will be the origin of a Maxwell's demon [40].

For an isothermal system, one can introduce Helmholtz's free energy function  $A = E - TS$ , then Eq. (49) becomes

$$dA = \mu dN - S dT - p dV + F_\alpha d\alpha. \quad (50)$$

And for an isothermal, isobaric information manipulation process without chemical reactions, one has Gibbs function  $G = E - TS + pV$  and  $dG = \mu dN - S dT + V dp + F_\alpha d\alpha = F_\alpha d\alpha$ . Note that while the first three terms contain "extensive" quantities  $N$ ,  $S$ , and  $V$ , the last term usually does not. It is nanothermodynamic [63]. Note also that for a feedback system that controls  $F_\alpha$ , one has  $\Theta = G - F_\alpha \alpha$  and  $d\Theta = -\alpha dF_\alpha$ .

Just as  $\mu$  is a function of temperature  $T$  in general, so is  $F_\alpha$ : It has an entropic part [64, 65]. This is where the "information" in Maxwell's demon enters thermodynamics. Eqs. (49) and (50), thus, are a grander First Law which now includes feedback information as a part of the conservation [39] with "informatic energy"  $F_\alpha d\alpha$ , on a par with heat energy  $T dS$ , mechanical energy  $p dV$ , and Gibbs' chemical energy  $\mu dN$ . Eq. (49) is the theory of absolute information in connection to controlling  $\alpha$ .

In engineering and biological research on complex systems, however, the notion of information often has a more subjective meaning, or meanings, usually hidden in the form of a statistical prior [66, 67]. One of the best examples, perhaps, is in current cellular biology: Many key biochemical processes inside a living cell are said to be "carrying out cellular signal transduction". Various biochemical activities and changing molecular concentrations are "interpreted" as "intracellular signals" that instruct a cell to respond to its environment. Here, two very different, but complementary, mathematical theories are equally valid: Since nearly all cellular biochemical reactions can be considered at constant temperature and volume, one describes the stochastic biochemical dynamics in terms of

Gibbs’ theory based on the  $\mu$  in Eq. (50). On the other hand, the same stochastic biochemical dynamics described in term of the probability theory can also be represented as an information processing machine with communication channels and transmissions of bits of information, carrying out a myriad of biological functions such as sensing, proofreading, timing, adaptation, and amplifications of signal magnitude, detection sensitivity, and response specificity [48]. The information flow narratives provide bioscientists a higher level of abstraction of a physicochemical reality [49].

Such an interpretive information theory, however, will lack the fundamental character of Eqs. (49) and (50). Still, as a multi-scale, coarse grained theory, some inequalities can be established [50, 51]. It is also noted that changing  $\alpha$  can always be mechanistically further represented in terms of changing geometric quantities such as volume and particle numbers via chemical reactions: The ultimate physical bases of information and its manipulation have to be matters and known forces.

We believe this dual possibility has a fundamental reason, rooted in Kolmogorov’s rigorous theory of probability: A probability space is an abstract object associated with which many different random variables, as measurements, are possible. At this point, it is interesting to read the preface of [46] written by E. T. Jaynes, who is considered by many as one of the greatest information theorists since Shannon: “From many years of experience with its applications in hundreds of real problems, our views on the foundations of probability theory have evolved into something quite complex, which cannot be described in any such simplistic terms as ‘pro-this’ or ‘anti-that’. For example, our system of probability could hardly be more different from that of Kolmogorov, in style, philosophy, and purpose. What we consider to be fully half of probability theory as it is needed in current applications — the principles for assigning probabilities by logical analysis of incomplete information — is not present at all in the Kolmogorov system.”

Then in an amazing candidness, Jaynes goes on: “Yet, when all is said and done, we find ourselves, to our own surprise, in agreement with Kolmogorov and in disagreement with its critics, on nearly all technical issues.”

**Acknowledgements.** We thank Professors P. Ao, M. Esposito, and H. Ge for helpful advices, Ying Tang, Lowell Thompson, Yue Wang, and Felix Ye for many discussions.

## A Derivation of work equalities

### A.1 The Jarzynski equality

The basic idea for the derivation is as follows: We first represent a path  $\mathbf{X}(t)$  by a discrete version with  $N$  steps and write the path probability in terms of the product of  $N$  transition probabilities given by the  $\prod_{i=0}^N(\cdots)$  in Eq. (51). Then the mean-exponential of negative

work

$$\left\langle e^{-\frac{1}{\epsilon^2} W[\mathbf{X}(\tau), \alpha(\tau)]} \right\rangle_{[\mathbf{X}(\tau), \alpha(\tau)]}$$

is [24]:

$$\begin{aligned} & \left\langle e^{-\frac{1}{\epsilon^2} W[\mathbf{X}_0, \dots, \mathbf{X}_N; \alpha_0, \dots, \alpha_N]} \right\rangle_{[\mathbf{X}_0, \dots, \mathbf{X}_N; \alpha_0, \dots, \alpha_N]} \\ &= \int \cdots \int \prod_{i=0}^N d\mathbf{X}_i \exp \left( -\frac{1}{\epsilon^2} \sum_{i=1}^N \left( \varphi(\mathbf{X}_i; \alpha_i) - \varphi(\mathbf{X}_i; \alpha_{i-1}) \right) \right) \\ & \quad \times \prod_{i=1}^N P(\mathbf{X}_i | \mathbf{X}_{i-1}; \alpha_{i-1}) p(\mathbf{X}_0, t_0; \alpha_0), \end{aligned} \quad (51)$$

in which the work from state  $(\mathbf{X}_i; \alpha_i)$  to state  $(\mathbf{X}_{i+1}; \alpha_{i+1})$  is defined as the difference in the global  $\varphi(\mathbf{x}; \alpha)$  with a common zero reference. This is a consequence of the First Law of Thermodynamics. Since equilibrium is attained at  $t_0$ ,  $p(\mathbf{x}_0; t_0) = f_{\mathbf{X}}^{eq}(\mathbf{x}_0; \alpha_0)$ . With the global  $\varphi(\mathbf{x}; \alpha) = -\epsilon^2 \ln f_{\mathbf{X}}^{eq}(\mathbf{x}; \alpha) - \epsilon^2 \ln Z(\alpha)$ , we have:

$$\begin{aligned} & \left\langle e^{-\frac{1}{\epsilon^2} W[\mathbf{X}_0, \dots, \mathbf{X}_N; \alpha_0, \dots, \alpha_N]} \right\rangle_{[\mathbf{X}_N, \dots, \mathbf{X}_0; \alpha_N, \dots, \alpha_0]} \\ &= \int \cdots \int \prod_{i=0}^N d\mathbf{X}_i \prod_{i=1}^N \frac{f_{\mathbf{X}}^{eq}(\mathbf{X}_i; \alpha_i) Z(\alpha_i)}{f_{\mathbf{X}}^{eq}(\mathbf{X}_i; \alpha_{i-1}) Z(\alpha_{i-1})} \cdot \prod_{i=1}^N P(\mathbf{X}_i | \mathbf{X}_{i-1}; \alpha_{i-1}) f_{\mathbf{X}}^{eq}(\mathbf{X}_0; \alpha_0) \\ &= \frac{Z(\alpha_n)}{Z(\alpha_0)} \int \cdots \int \prod_{i=0}^N d\mathbf{X}_i \prod_{i=1}^N P(\mathbf{X}_i | \mathbf{X}_{i-1}; \alpha_{i-1}) \prod_{i=1}^N f_{\mathbf{X}}^{eq}(\mathbf{X}_{i-1}; \alpha_{i-1}) \bigg/ \prod_{i=1}^N f_{\mathbf{X}}^{eq}(\mathbf{X}_i; \alpha_{i-1}) \\ &= \frac{Z(\alpha_n)}{Z(\alpha_0)}. \end{aligned} \quad (52)$$

Since we have defined in Sec. 3.1 the free energy as:  $A_2(\alpha) = -\epsilon^2 \ln Z(\alpha)$ , thus we obtain the Jarzynski equality:

$$\left\langle e^{-\frac{1}{\epsilon^2} W[\mathbf{X}(\tau), \alpha(\tau)]} \right\rangle_{[\mathbf{X}(\tau), \alpha(\tau)]} = e^{-\frac{1}{\epsilon^2} \Delta A_2}. \quad (53)$$

In a very similar vein, for the macroscopic thermodynamic variables  $(E, \tilde{\alpha})$ , one defines the work done to the system by the external environment through controlling  $\tilde{\alpha}(t)$  with rate  $\dot{\tilde{\alpha}}$ :

$$W[E(\tau), \tilde{\alpha}(\tau)] = - \int_0^t \tilde{F}_{\tilde{\alpha}}(E, \tilde{\alpha}) \dot{\tilde{\alpha}} d\tau. \quad (54)$$

Write  $\varphi_\tau(\tilde{\alpha}) = E(\tau, \tilde{\alpha})$ . Then the discretized  $\left\langle e^{-\frac{1}{\epsilon^2} W[E(\tau), \tilde{\alpha}(\tau)]} \right\rangle_{[E(\tau), \tilde{\alpha}(\tau)]}$  is:

$$\begin{aligned} & \left\langle e^{-\frac{1}{\epsilon^2} W[E_0, \dots, E_N; \tilde{\alpha}_0, \dots, \tilde{\alpha}_N]} \right\rangle_{[E_0, \dots, E_N; \tilde{\alpha}_0, \dots, \tilde{\alpha}_N]} \\ &= \int \cdots \int \prod_{i=0}^N dE_i \prod_{i=1}^N P(E_i | E_{i-1}; \tilde{\alpha}_{i-1}) p(E_0, t_0; \tilde{\alpha}_0) \times \\ & \exp \left( - \sum_{i=1}^N \frac{\varphi_i(\tilde{\alpha}_i) - \varphi_i(\tilde{\alpha}_{i-1})}{\epsilon^2} + S(E_i, \tilde{\alpha}_i) - S(E_i, \tilde{\alpha}_{i-1}) \right), \end{aligned} \quad (55)$$

where  $S(E, \tilde{\alpha})$  is defined in Eq. (43). On the other hand, equilibrium probability density function of  $E$  at  $(E_i, \tilde{\alpha}_i)$  is:

$$\begin{aligned} f_{\mathbf{E}}^{eq}(E_i, \tilde{\alpha}_i) &= \frac{1}{Z(\tilde{\alpha}_i)} \oint_{\varphi(\mathbf{x}; \tilde{\alpha})=E_i} e^{-\varphi(\mathbf{x}; \tilde{\alpha}_i)/\epsilon^2} \frac{d\Sigma^{n-1}}{||\nabla_{\mathbf{x}} \varphi(\mathbf{x}; \tilde{\alpha}_i)||} \\ &= Z^{-1}(\tilde{\alpha}_i) \exp \left( - \frac{\varphi_i(\tilde{\alpha}_i)}{\epsilon^2} + S(E_i, \tilde{\alpha}_i) \right). \end{aligned} \quad (56)$$

Hence, we have

$$\begin{aligned} & \left\langle e^{-\frac{1}{\epsilon^2} W[E_0, \dots, E_N; \tilde{\alpha}_0, \dots, \tilde{\alpha}_N]} \right\rangle_{[E_0, \dots, E_N; \tilde{\alpha}_0, \dots, \tilde{\alpha}_N]} \\ &= \int \cdots \int \prod_{i=0}^N dE_i \prod_{i=1}^N \frac{f_{\mathbf{E}}^{eq}(E_i, \tilde{\alpha}_i) Z(\tilde{\alpha}_i)}{f_{\mathbf{E}}^{eq}(E_i, \tilde{\alpha}_{i-1}) Z(\tilde{\alpha}_{i-1})} \left( \prod_{i=1}^N P(E_i | E_{i-1}, \tilde{\alpha}_{i-1}) f_{\mathbf{E}}^{eq}(E_0, \tilde{\alpha}_0) \right) \\ &= \frac{Z(\tilde{\alpha}_n)}{Z(\tilde{\alpha}_0)} \int \cdots \int \prod_{i=0}^N dE_i \prod_{i=1}^N P(E_i | E_{i-1}, \tilde{\alpha}_{i-1}) \prod_{i=1}^N f_{\mathbf{E}}^{eq}(E_{i-1}, \tilde{\alpha}_{i-1}) \Big/ \prod_{i=1}^N f_{\mathbf{E}}^{eq}(E_i, \tilde{\alpha}_{i-1}) \\ &= \frac{Z(\tilde{\alpha}_n)}{Z(\tilde{\alpha}_0)} = e^{-\frac{1}{\epsilon^2} \Delta A_2(\alpha)}. \end{aligned} \quad (57)$$

Therefore, the log-mean exponential of minus work is equal to the minus of free energy difference.

## A.2 Crooks' approach

Instead of introducing stochastic work functional, G. E. Crooks' approach recognizes the important role of *time reversal* trajectory  $\tilde{\mathbf{X}}(t)$ , and the deep relationship between work, energy, and dissipation, e.g., entropy production. Let us consider the path probability of a backward trajectory  $\mathcal{P}[\tilde{\mathbf{X}}(\tau) | \tilde{\mathbf{X}}(0); \tilde{\alpha}(\tau)]$  against a forward one  $\mathcal{P}[\mathbf{X}(\tau) | \mathbf{X}(0); \alpha(\tau)]$ , in which  $(\tilde{\mathbf{X}}(\tau); \tilde{\alpha}(\tau)) = (\mathbf{X}(t - \tau); \alpha(t - \tau))$ , where the initial and final distribution of  $\mathbf{X}(\tau)$  follow the equilibrium distribution. We solve for  $\mathcal{P}[\mathbf{X}(\tau) | \mathbf{X}(0); \alpha(\tau)]$  from the probability of a Brownian motion whose increments are multivariate Gaussian:

$$\frac{1}{\epsilon} \Gamma(\alpha_i)^{-1} \left( \mathbf{X}_{i+1} - \mathbf{X}_i - M(\alpha_i) \mathbf{X}_i \Delta \tau \right) = \mathbf{B}_{t_{i+1}} - \mathbf{B}_{t_i}. \quad (58)$$

Hence, the probability density functional of a path  $[\mathbf{X}_0, \dots, \mathbf{X}_N | \mathbf{X}_0; \alpha_0, \dots, \alpha_N]$  is:

$$\begin{aligned} \mathcal{P}[\mathbf{X}_0, \dots, \mathbf{X}_N | \mathbf{X}_0; \alpha_0, \dots, \alpha_N] &= \prod_{i=0}^{N-1} P(\mathbf{X}_{i+1} | \mathbf{X}_i; \alpha_i) \\ &= \prod_{i=0}^{N-1} \frac{1}{(\pi \Delta \tau)^{n/2}} e^{-\frac{1}{\epsilon^2 \Delta \tau} (\Gamma(\alpha_i)^{-1} (\mathbf{X}_{i+1} - \mathbf{X}_i) - \Gamma(\alpha_i)^{-1} M(\alpha_i) \mathbf{X}_i \Delta \tau)^2}. \end{aligned} \quad (59)$$

Here  $(\mathbf{v}(\mathbf{X}, \alpha))^2 \equiv (\mathbf{v}(\mathbf{X}, \alpha))^T (\mathbf{v}(\mathbf{X}, \alpha))$ . Then the probability density functional of the inverse path  $[\mathbf{X}_N, \dots, \mathbf{X}_0 | \mathbf{X}_N; \alpha_N, \dots, \alpha_0]$  is:

$$\begin{aligned} \mathcal{P}[\mathbf{X}_N, \dots, \mathbf{X}_0 | \mathbf{X}_N; \alpha_N, \dots, \alpha_0] &= \prod_{i=1}^N P(\mathbf{X}_{i-1} | \mathbf{X}_i; \alpha_i) \\ &= \prod_{i=1}^N \frac{1}{(\pi \Delta \tau)^{n/2}} e^{-\frac{1}{\epsilon^2 \Delta \tau} (\Gamma(\alpha_i)^{-1} (\mathbf{X}_{i-1} - \mathbf{X}_i) - \Gamma(\alpha_i)^{-1} M(\alpha_i) \mathbf{X}_i \Delta \tau)^2}. \end{aligned} \quad (60)$$

Therefore, offsetting by a normalization factor, an infinite-dimensional functional integral,

$$\begin{aligned} \mathcal{P}[\mathbf{X}(\tau) | \mathbf{X}(0); \alpha(\tau)] &\propto \exp \left[ -\frac{1}{\epsilon^2} \int_0^t \left( \Gamma(\alpha)^{-1} d\mathbf{X} / \sqrt{d\tau} - \Gamma(\alpha)^{-1} M(\alpha) \mathbf{X} \sqrt{d\tau} \right)^2 \right] \\ &= \exp \left[ -\frac{1}{\epsilon^2} \int_0^t \left( \Gamma(\alpha)^{-1} \dot{\mathbf{X}} - \Gamma(\alpha)^{-1} M(\alpha) \mathbf{X} \right)^2 d\tau \right]. \end{aligned} \quad (61)$$

Probability of the backward path can be found by substituting  $\tau$  with  $t - \tau$ :

$$\mathcal{P}[\check{\mathbf{X}}(\tau) | \check{\mathbf{X}}(0); \check{\alpha}(\tau)] \propto \exp \left[ -\frac{1}{\epsilon^2} \int_0^t \left( -\Gamma(\alpha)^{-1} \dot{\mathbf{X}} - \Gamma(\alpha)^{-1} M(\alpha) \mathbf{X} \right)^2 d\tau \right]. \quad (62)$$

Therefore, we have an equality for heat dissipation:

$$\begin{aligned} &\frac{\mathcal{P}[\check{\mathbf{X}}(\tau) | \check{\mathbf{X}}(0); \check{\alpha}(\tau)]}{\mathcal{P}[\mathbf{X}(\tau) | \mathbf{X}(0); \alpha(\tau)]} \\ &= \exp \left[ \frac{2}{\epsilon^2} \int_0^t \left( \dot{\mathbf{X}}^T (\Gamma(\alpha) \Gamma(\alpha)^T)^{-1} M(\alpha) \mathbf{X} + \mathbf{X}^T M(\alpha)^T (\Gamma(\alpha) \Gamma(\alpha)^T)^{-1} \dot{\mathbf{X}} \right) d\tau \right] \\ &= \exp \left[ \frac{2}{\epsilon^2} \int_0^t \left( \dot{\mathbf{X}}^T D^{-1}(\alpha) M(\alpha) \mathbf{X} \right) d\tau \right] \\ &= \exp \left[ \frac{2}{\epsilon^2} \int_0^t \dot{\mathbf{X}}^T \Xi^{-1}(\alpha) \mathbf{X} d\tau + \frac{2}{\epsilon^2} \int_0^t \dot{\mathbf{X}}^T (D^{-1}(\alpha) M(\alpha) - \Xi^{-1}(\alpha)) \mathbf{X} d\tau \right] \\ &= \exp \left\{ \frac{Q[\mathbf{X}(\tau), \alpha(\tau)]}{\epsilon^2} + \frac{Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)]}{\epsilon^2} \right\}. \end{aligned} \quad (63)$$

Hatano and Sasa, following Oono and Paniconi, called the term

$$Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)] = -2 \int_0^t \dot{\mathbf{X}}^T (D^{-1}(\alpha) M(\alpha) - \Xi^{-1}(\alpha)) \mathbf{X} d\tau \quad (64)$$

house-keeping heat [24].

Since we start and end with equilibrium distributions with the corresponding  $\tilde{\alpha}$ ,

$$\begin{aligned} p(\mathbf{X}_0; \alpha_0) &= f_{\mathbf{X}}^{eq}(\mathbf{X}_0; \alpha_0) = \frac{1}{Z(\alpha_0)} \exp \left[ -\frac{\mathbf{X}_0^T U(\alpha_0) \mathbf{X}_0}{\epsilon^2} \right]; \\ p(\mathbf{X}_N; \alpha_N) &= f_{\mathbf{X}}^{eq}(\mathbf{X}_N; \alpha_N) = \frac{1}{Z(\alpha_N)} \exp \left[ -\frac{\mathbf{X}_N^T U(\alpha_N) \mathbf{X}_N}{\epsilon^2} \right]. \end{aligned} \quad (65)$$

Therefore,

$$\begin{aligned} \frac{\mathcal{P}[\tilde{\mathbf{X}}(\tau), \tilde{\alpha}(\tau)]}{\mathcal{P}[\mathbf{X}(\tau), \alpha(\tau)]} &= \frac{\mathcal{P}[\tilde{\mathbf{X}}(\tau) | \tilde{\mathbf{X}}(0); \tilde{\alpha}(\tau)] p(\mathbf{X}_N; \alpha_N)}{\mathcal{P}[\mathbf{X}(\tau) | \mathbf{X}(0); \alpha(\tau)] p(\mathbf{X}_0; \alpha_0)} \\ &= \exp \left( \frac{Q[\mathbf{X}(\tau), \alpha(\tau)] - Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)] - \Delta\varphi + \Delta A_2}{\epsilon^2} \right) \\ &= \exp \left( \frac{-W[\mathbf{X}(\tau), \alpha(\tau)] - Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)] + \Delta A_2}{\epsilon^2} \right). \end{aligned} \quad (66)$$

Now taking ensemble average of the trajectories  $[\mathbf{X}(\tau), \alpha(\tau)]$  over  $\frac{\mathcal{P}[\tilde{\mathbf{X}}(\tau), \tilde{\alpha}(\tau)]}{\mathcal{P}[\mathbf{X}(\tau), \alpha(\tau)]}$  gives:

$$\begin{aligned} \int \mathcal{D}[\mathbf{X}(\tau), \alpha(\tau)] \mathcal{P}[\tilde{\mathbf{X}}(\tau), \tilde{\alpha}(\tau)] &= \left\langle \frac{\mathcal{P}[\tilde{\mathbf{X}}(\tau), \tilde{\alpha}(\tau)]}{\mathcal{P}[\mathbf{X}(\tau), \alpha(\tau)]} \right\rangle_{[\mathbf{X}(\tau), \alpha(\tau)]} \\ &= e^{\Delta A_2} \left\langle e^{-W[\mathbf{X}(\tau), \alpha(\tau)] - Q_{hk}[\mathbf{X}(\tau), \alpha(\tau)]} \right\rangle_{[\mathbf{X}(\tau), \alpha(\tau)]}. \end{aligned} \quad (67)$$

When the system is in detailed balance, Crooks' approach recovers the Jarzynski equality. If one chooses the global energy  $\varphi$  with zero reference for each own equilibrium, i.e.,  $\Delta A_2 = 0$  for all  $\alpha$ , then it recovers the Hatano-Sasa equality.

## References

- [1] Landau L D and Lifshitz E M 1980 *Statistical Physics* 3rd ed Part 1 (*Course of Theoretical Physics vol 5*) Sykes J B and Kearsley M J (transl) (Oxford, UK: Pergamon Press)
- [2] Callen H B 1985 *Thermodynamics and an Introduction to Thermostatistics* 2nd ed (New York: Wiley)
- [3] Onsager L and Machlup S 1953 *Phys. Rev.* **91** 1505–12
- [4] Machlup S and Onsager L 1953 *Phys. Rev.* **91** 1512–15
- [5] Wax N 1954 *Selected Papers on Noise and Stochastic Processes* (New York: Dover)
- [6] Cox R T 1952 *Rev. Mod. Phys.* **24** 312–20

- [7] Lax M 1960 *Rev. Mod. Phys.* **32** 25–64
- [8] Lax M 1966 *Rev. Mod. Phys.* **38** 359–79
- [9] Fox R R 1978 *Phys. Rep.* **48** 180–283
- [10] Qian H 2001 *Proc. Roy. Soc. A: Math. Phys. Engr. Sci.* **457** 1645–55
- [11] Kwon C, Ao P and Thouless D J 2005 *Proc. Natl. Acad. Sci. U. S. A.* **102** 13029–33
- [12] Dotsenko V, Macioek A, Vasilyev O and Oshanin G 2013 *Phys. Rev. E* **87** 062130
- [13] Seifert U 2012 *Rep. Prog. Phys.* **75** 126001
- [14] Ge H 2014 *Adv. Math.* (in China) **43** 161–74
- [15] van den Broeck C and Esposito M 2015 *Physica A* **418** 6–16
- [16] Altaner B 2014 Foundations of stochastic thermodynamics (arXiv:1410.3983)
- [17] Qian M-P, Qian M and Gong G-L 1991 *Contemp. Math.* **118** 255–61
- [18] Seifert U 2005 *Phys. Rev. Lett.* **95** 040602
- [19] Ge H and Jiang D-Q 2007 *J. Phys. A: Math. Theor.* **40** F713-23
- [20] Jarzynski C 1997 *Phys. Rev. Lett.* **78** 2690–3
- [21] Kurchan J 1998 *J. Phys. A: Math. Gen.* **31** 3719-29
- [22] Crooks G 1999 *Phys. Rev. E.* **60** 2721–6
- [23] Lebowitz J L and Spohn H 1999 *J. Stat. Phys.* **95** 333–65
- [24] Hatano T and Sasa S-I 2001 *Phys. Rev. Lett.* **86** 3463–6
- [25] Jiang, D-Q Qian M and Qian M-P 2004 *Mathematical Theory of Nonequilibrium Steady States (Lect. Notes Math.)* vol 1833 (New York: Springer)
- [26] Zhang X-J, Qian H and Qian M 2012 *Phys. Rep.* **510** 1–86
- [27] Ge H, Qian M and Qian H 2012 *Phys. Rep.* **510** 87–118
- [28] Lamb J S W and Roberts J A G 1998 *Physica D* **112** 1–39
- [29] van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* 2nd ed (Amsterdam: North Holland)



- [30] Risken H 1996 *The Fokker-Planck Equation, Methods of Solution and Applications* (Berlin: Springer)
- [31] Ao P 2004 *J. Phys. A: Math. Gen.* **37** L25–30
- [32] Qian H 2014 *Phys. Lett. A* **378** 609–16
- [33] Qian H 2014 Thermodynamics of the general diffusion process: Equilibrium super-current and nonequilibrium driven circulation with dissipation (arXiv:1412.5925)
- [34] Wang J, Xu L and Wang E 2008 *Proc. Natl. Acad. Sci. U. S. A.* **105** 12271–6
- [35] Ma Y-A, Tan Q, Yuan R, Yuan B and Ao P 2014 *Int. J. Bifur. Chaos* **24** 1450015
- [36] Pfund D M, Lee L L and Cochran H D 1990 *Int. J. Thermophys.* **11** 73–86
- [37] Noyes R M 1961 *J. Chem. Phys.* **34** 1983–5
- [38] Ge H and Qian H 2011 Maximum entropy principle, equal probability *a priori*, and Gibbs paradox (arXiv:1105.4118)
- [39] Toyabe S, Sagawa T, Ueda M, Muneyuki E and Sano M 2010 *Nature Phys.* **6** 988–92
- [40] Mandal D and Jarzynski C 2012 *Proc. Natl. Acad. Sci. U. S. A.* **109** 11641–5
- [41] Khinchin A Ya 1949 *Mathematical Foundations of Statistical Mechanics* (New York: Dover)
- [42] Sekimoto K 2010 *Stochastic Energetics (Lect. Notes in Physics)* vol 799 (New York: Springer)
- [43] Tang Y, Yuan R, Ao P 2014 Work relations connecting nonequilibrium steady states without detailed balance (arXiv:1409.0498)
- [44] Kwon C, Noh J D and Park H 2013 *Phys. Rev. E* **88** 062102
- [45] Ma Y-A, Qian H and Ye F X-F 2015 On entropy distribution of high-dimensional random variables: Uncertainty quantification and Helmholtz-Boltzmann statistics (Manuscript in preparation)
- [46] Jaynes E T 2003 *Probability Theory: The Logic of Science* (Cambridge, UK: Univ. Press)
- [47] Pressé S, Ghosh K, Lee J and Dill K A 2013 *Rev. Mod. Phys.* **85** 1115–41
- [48] Hartwell L H, Hopfield J J, Leibler S and Murray A W 1999 *Nature* **402** C47–52

- [49] Qian H and Roy S 2012 *IEEE Trans. NanoBiosci.* **11** 289–95
- [50] Santillán M and Qian H 2011 *Phys. Rev. E* **83** 041130
- [51] Esposito M 2012 *Phys. Rev. E* **85** 041125
- [52] Anderson P W 1972 *Science* **177** 393–6
- [53] Ao P, Qian H, Tu Y and Wang J 2013 A theory of mesoscopic phenomena: Time scales, emergent unpredictability, symmetry breaking and dynamics across different levels (arXiv:1310.5585)
- [54] Kolmogorov A N 1931 *Mathematische Annalen* **104** 415–58
- [55] Wegscheider R 1902 *Zeitschrift für Physikalische Chemie* **39** 257–303
- [56] Lewis G N 1925 *Proc. Natl. Acad. Sci. U. S. A.* **11** 179–83
- [57] Onsager L 1931 *Phys. Rev.* **37** 405–26
- [58] Prigogine I 1947 *Etude thermodynamique des phénomènes irréversibles* (Liege: Desoer)
- [59] Tolman R C and Fine P C 1948 *Rev. Mod. Phys.* **20** 51–77
- [60] Ma Y-A and Qian H 2014 The Helmholtz theorem for the Lotka-Volterra equation, the extended conservation relation, and stochastic predator-prey dynamics (arXiv:1405.4311)
- [61] Fogedby H C and Imparato A 2012 *J. Stat. Mech. Theor. Exp.* P04005
- [62] Kolmogorov A N 1968 *Int. J. Comput. Math.* **2** 157–68
- [63] Hill T L 1994 *Thermodynamics of Small Systems* (New York: Dover)
- [64] Qian H and Hopfield J J 1996 *J. Chem. Phys.* **105** 9292–6
- [65] Qian H 2002 *Phys. Rev. E* **64**, 022101
- [66] Polettini M 2012 *EPL* **97** 30003
- [67] Polettini M 2013 Of dice and men. Subjective priors, gauge invariance, and nonequilibrium thermodynamics (arXiv:1307.2057)